ALMA ACA Spectrometer
New GPU-based Spectrometer for Total Power Array

〇Takashi Nakamoto, Manabu Watanabe, Scott Zang, Makoto Shizugami,
Tetsuhiro Minamidani, George Kosugi (NAOJ), Jongsoo Kim (KASI)

Abstract

Korea Astronomy and Space Science Institute (KASI) and National Astronomical Observatory of Japan (NAOJ) are jointly developing the Atacama Compact Array (ACA) Spectrometer for the Atacama Large Millimeter/submillimeter Array (ALMA). This is a new spectrometer that can process the signals from four total power (TP) antennas using graphical processing units (GPUs). The development team achieved a significant milestone, the Critical Design and Manufacturing Review, in December 2019. This report presents key design elements of the ACA Spectrometer.

1. Background

The Total Power (TP) Array comprises four 12 m antennas and is a subset of the Atacama Large Millimeter/submillimeter Array (ALMA). While the other ALMA antennas are used for interferometric observations, the TP Array is responsible for single-dish observations. Auto-correlation outputs from the array enable observers to recover large-scale emissions from extended science targets that may have been spatially filtered out of interferometric observations.

The TP Array signals are currently processed by the ACA Correlator, an FX-type correlator, which also processes the signals from twelve 7 m antennas. The correlator is equipped with high-speed Field Programmable Gate Arrays (FPGAs) for performing 16 bit fixed-point fast Fourier transforms (FFTs, the F part) and multiplications (the X part). A re-quantization step after the FFT converts the 16-bit output into 4 bits. It is necessary to bring the data rate and the number of multiplication calculations down to a reasonable range, but it adds 1.1% to the sensitivity loss [1]. Additionally, the limited depth of the fixed-point calculations can cause an overflow during FFT computations. More importantly, observations of the TP Array and the 7 m Array using one correlator have restricted science operations.

A new spectrometer for the TP Array, the ACA Spectrometer, is currently under development to mitigate those problems. It performs 32 bit floating-point calculations in both the F and X parts using GPUs to reduce the sensitivity loss and practically eliminate the risk of overflow. It also produces both auto- and cross-correlation outputs. Its cross-correlation capability enables calibration observations (e.g., pointing) of the TP Array independently of the 7 m Array.

2. Input Data Flow

For all ALMA antennas, including those of the TP Array, the radio frequency signal for each polarization is first downconverted to the intermediate frequency signal of 8 GHz bandwidth. It is then divided into four 2 GHz baseband signals, which are digitized by 4 Gsps analog-digital converters with 3 bit resolution (the B, C, and D bits). This results in a data rate of 96 Gbps (2 polarizations, 4 basebands, 4 Gsps, 3 bits) in total from each antenna. The digital signals of two polarizations in the same baseband are combined to make a baseband pair (BBpr) signal and transmitted over three optical fibers. Each fiber carries an 8 Gbps stream of B, C, or D bits using a dedicated protocol [2]. The digital signals over the optical fibers from the TP Array are split so that the ACA Spectrometer can simultaneously receive the same signals as the ACA Correlator.

The ACA Spectrometer system has four ACA Spectrometer Modules (ASMs), each receiving the BBpr signals of the same baseband from all four TP antennas. The ASM comprises dual CPUs (Intel Xeon Gold 6144 or 6244), twelve DDR4-2666 memory modules (six modules per CPU), four GPUs (Nvidia TITAN V) and two data transmission system receivers with PCI Express bus interface (DRXPs) on a Supermicro 4029GP-TRT server running Red Hat Enterprise Linux 7. The DRXP is a custom-made board with a Xilinx UltraScale+ FPGA that receives, decodes, and transfers signals from two antennas to the system memory via PCI Express 3.0 ×16 lanes (PCIe3 ×16) using direct memory access (DMA). The DRXP sends the data block by block to maximize the bandwidth. Each block is 4 MiB, which is the largest contiguous physical memory block that the Linux memory management system can allocate at once. Calculations are assigned to the GPUs in a round-robin manner. Contiguous 16 ms data (equivalent to 11.4 blocks of data) from all four antennas are sent to a GPU to complete the FX calculations within one GPU (Figure 1). With this data flow, each GPU processes 8 Gsps (3 GB/s) of data on average.

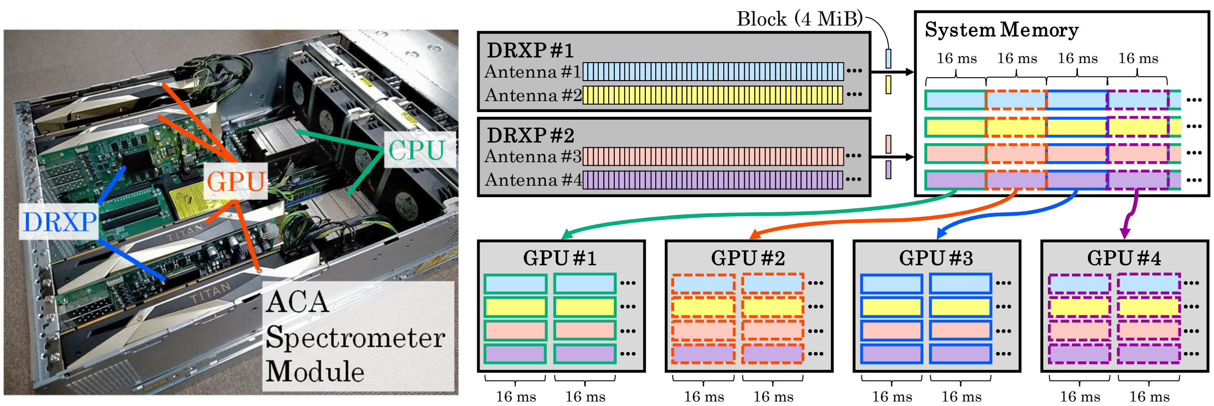

Figure 1. ASM with top cover removed (left) and a schematic of the data streams from four antennas split and sent to four GPUs (right).

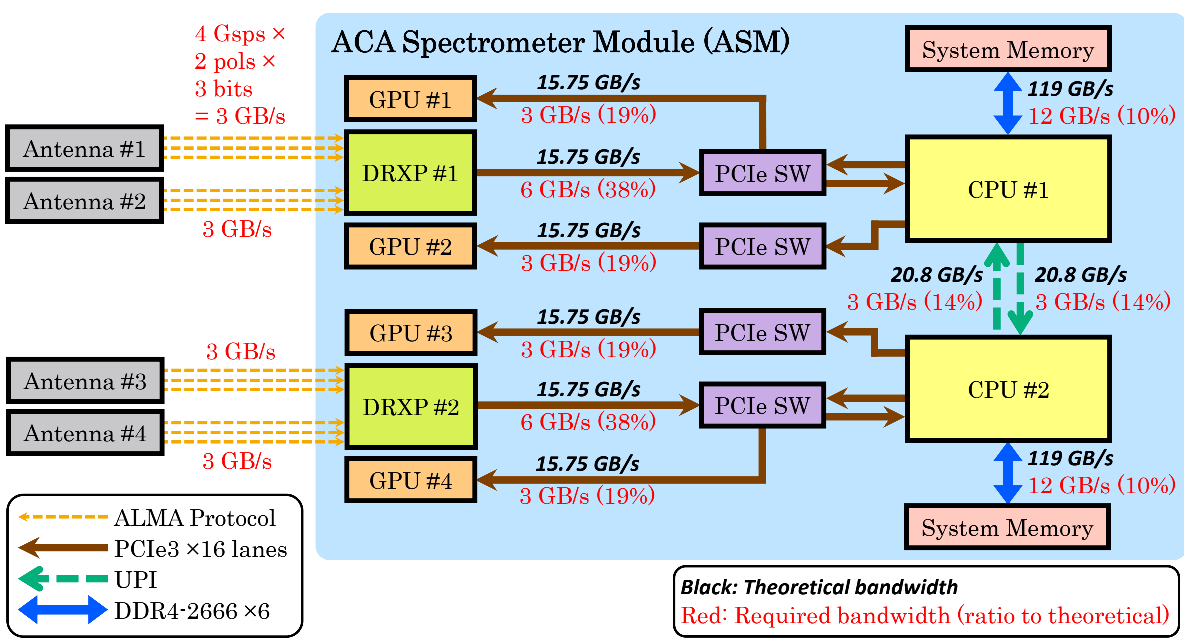

Figure 2. Theoretical and required bandwidths of communication paths from four antennas to four GPUs.

A Linux kernel module that was developed for the DRXP allocates DMA buffers in the system memory. Our application program written in C++ calls `cudaHostRegister()` with the `cudaHostRegisterIoMemory` flag for each block to share the DMA buffers between the DRXPs and the GPUs. GPU programs in Compute Unified Device Architecture (CUDA) see such memory areas as pinned memory. The theoretical maximum bandwidth of PCIe3 ×16 is 15.75 GB/s. Our benchmark test shows that the data transfer from the system memory to a GPU using the pinned memory can achieve a bandwidth of approximately 12 GB/s, four times more than the required 3 GB/s. Another benchmark confirms that the same bandwidth can be achieved when sending data to all four GPUs simultaneously.

3. GPU Performance

The GPUs convert 3 bit integers into 32 bit floating-point values followed by FFT using `cufftExecR2C`. The ACA Spectrometer internally performs 2^{20} points FFTs for the finest spectral resolution of 3.815 kHz. This is the most expensive computation task in the ACA Spectrometer. Due to a mismatch between the ALMA integration time granularities (1 ms = 4×10^6 samples) and the number of FFT points, the GPU runs four 2^{20}-points FFTs with 4.6% overlap between FFT segments for every 1 ms of sampling. This maintains compatibility with the existing ACA Correlator [1]. Our benchmark results in Figure 3 show that the overlap significantly degrades the FFT performance, which is thought to be caused by non-sequential memory access. However, the benchmark result still proves that TITAN V can process incoming data streams of 8 Gsps. Another benchmark test demonstrated that TITAN V could perform all major calculations in real time, including the bit conversions and the FX calculations.

The existing ALMA archive system limits the output data rate from the ACA Spectrometer. The ACA Spectrometer will share the archive system with the two correlators currently in use. The maximum data rate of 17 MB/s will be allocated to the ACA Spectrometer. Based on the observation configuration imposed by the observatory control system, the GPUs perform spectral averaging and 1 ms time integration, and the CPUs perform longer time integration so that the average output rate does not exceed the allocation.

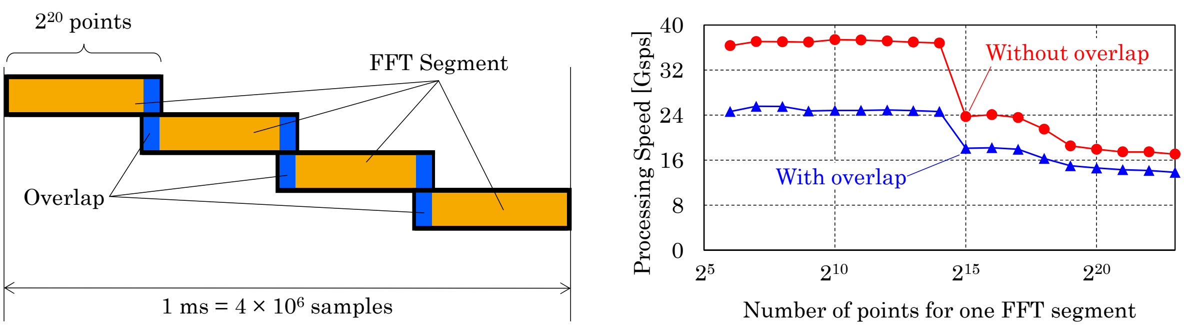

Figure 3. FFT segmentation with overlap for the shortest integration time, 1 ms (left), and `cufftExecR2C` benchmark result of one TITAN V for various FFT points with and without overlap (right).

4. GPU Cooling Efficiency at High Altitude

The ACA Spectrometer will be installed in the Array Operation Site Technical Building (AOS TB) at an altitude of about 5,000 m. Due to the thin air at this high altitude, there is a risk of GPU overheating even with cooling fans. Once the temperature goes over a threshold, the GPUs slow their clock frequency impairing calculation performance. The GPU temperature T follows Newton's law of cooling

$$Q = \bar{h}(T - T_\infty) \tag{1}$$

where Q is the GPU power consumption, T_∞ is the inlet air temperature and $\bar{h}$ is the average heat transfer coefficient. The average Nusselt number $\overline{Nu}$ of a heating surface parallel to laminar airflow is

$$\overline{Nu} = 0.664Re^{\frac{1}{2}}Pr^{\frac{1}{3}} \tag{2}$$

where Re is the Reynolds number and Pr is the Prandtl number [3]. In general, dynamic viscosity and thermal conductivity depend on pressure. However, the dependency on pressure is negligible from hundreds to hundreds of thousands of Pascals. Therefore, we obtain this practical relation

$$\bar{h} \propto \sqrt{\rho v}. \tag{3}$$

where ρ is the air density and v is the flow velocity.

The legacy experimental data of the ACA Correlator confirms this relation. One module of the correlator was placed in a depressurized chamber. The surface temperature of FPGAs, memory modules and other electrical components were measured under two conditions: (i) 28.7 °C, 760 mmHg (equivalent to sea level); and (ii) 30.1 °C, 409 mmHg (equivalent to an altitude of 5,000 m). The density of dry air in these conditions is 1.169 and 0.626 kg/m^3, respectively. Therefore, the heat transfer coefficient ratio of condition (i) to condition (ii) must be 1.37 according to relation (3). The mean heat transfer coefficient ratio derived from the temperature data at 20 different measurement points was 1.39 with a standard deviation of 0.038. This is close to 1.37, which proves relation (3).

At an early design phase of the ACA Spectrometer, we conducted a similar experiment using two TITAN Xps and two GeForce TITAN Xs, which are old-generation GPUs with appearance like TITAN V. They were tested with three different GPU fan speeds (60, 80, and 100%) in the AOS TB and NAOJ Mitaka campus at almost sea level. A dummy benchmark program was used to control the power consumption Q, which was recorded with the GPU temperature T via Nvidia management library (NVML). The least square method for equation (1) was applied to estimate the heat transfer coefficient $\bar{h}$. This experiment showed that the heat transfer coefficient ratios ranged from 1.43 to 1.67 with a standard deviation of 0.077; which is significantly higher than 1.37, but the reason for this is not known. Once TITAN V became available, the sea level heat transfer coefficients of four TITAN Vs were measured at the NAOJ Mitaka campus. The relation between the GPU temperature and power consumption at the AOS TB was derived using the worst heat transfer coefficient ratio (1.67) to be safe. The power consumption was estimated using a separate benchmark program, and it was finally concluded that the GPU temperature at the AOS TB would not exceed the threshold.

5. Summary and Future Work

The design of the ACA Spectrometer and hardware selection has been completed. The required software is currently under development, and the ACA Spectrometer is expected to be installed in 2021.

Owing to recent advances in GPU technology, almost all functions will be implemented in application-level software using C++ and CUDA. This architecture allows us to implement software independent of hardware. In the future, it will also enable us to support more observation configurations, add new features, and extend the performance (e.g., more antennas, more frequency channels, and wider bandwidth) with less effort than FPGA-based correlators.

6. References

1. Takeshi Kamazaki et al., Digital Spectro-Correlator System for the Atacama Compact Array of the Atacama Large Millimeter/Submillimeter Array, Publications of the Astronomical Society of Japan, Volume 64, Issue 2, 25 April 2012, 29

2. Robert W. Freund, Digital Transmission System Signaling Protocol, ALMA Memo No. 420 (September 6, 2002)

3. Jingquan Cheng, Forced Air Cooling at High Altitude, MMA Memo 203 (March 24, 1998)